

Supplemental Fig. 4. Distribution of intron start deviations from group averages for DEFLs in *Arabidopsis thaliana*. The percentage of genes having an intron start deviation (from the average of all members in its subgroup, taken relative to the translation start) of $x \pm 1.5$ bp is plotted for three categories of genes: expressed = genes whose intron boundaries are known from experimentally sequenced cDNAs in the literature (52 total); predicted = genes whose donor and acceptor were predicted by the NetPlantGene server, <http://www.cbs.dtu.dk/services/NetPGene/> (90 total); aligned = genes whose donor and/or acceptor were predicted by visual alignment with other homologs in the same subgroup (139 total). Genes or pseudogenes lacking an intron (see Supp. Table I) and second introns (1 case only) were not included. Results from the RT-PCR and sequencing carried out for this project are also not reflected in this plot.